NAME
     netfind - Internet white pages tool

SYNOPSIS
     netfind [-h] [-s] [-t] [-T#] [-D#] [-H#]
          [-m[file]]" "[-d<Aacfhlmnrst>] name key key ..."

DESCRIPTION
     Netfind is a utility that provides an Internet "white pages"
     directory facility.   It runs on SunOS 4.0 or more recent
     systems. Given the name of a person on the Internet and a
     rough description of where the person works, the tool
     attempts to locate electronic mailbox information about the
     person.   To make basic use of the tool, first set the
     environment variable SEED_DB to point to the inverted index
     file where the netfind seed database exists (INDEX). (You
     system administrator can tell you where this is.)   Then,
     enter a command such as

          netfind schwartz university colorado boulder

     The name argument specifies the person being sought by first
     name, last name, or login name. The keys describe where the
     person works. They could consist of any combination of
     strings describing the name of the institution, the
     city/state/country, or the type of institution (e.g., "edu",
     "com", "mil", etc.)   Keys are case insensitive and may be
     specified in any order. Using more than one key implies the
     logical AND of the keys. Because of this, specifying too
     many keys may actually cause searches to fail. If this hap-
     pens, try specifying fewer keys, e.g.,

          netfind schwartz boulder

     When netfind runs, it displays a trace of the parallel
     search progress, along with the results of the searches. If
     you wish to disable the added output (causing netfind to
     print only information from successful searches, plus warn-
     ings if no responses came back and there were potentially
     transient host or network failures), use

          netfind -dfmn ...

     See the BASIC MECHANISM section for a discussion of how
     searches proceed, and the OPTIONS section for a more
     detailed discussion of these and other options.

BASIC MECHANISM
     While netfind can easily be used without understanding how
     searches proceed, a basic understanding will help you make
     better use of the tool. When a search is requested, netfind

consults a "seed" database to find potential machines to
search, using the specified keys. These and other machines
can then be searched concurrently using the finger protocol
and SMTP, the electronic mail protocol.

In order to reduce Internet loading, netfind places limita-
tions on the number of administrative domains to be searched
(3) and the number of machines fingered (50). If more than
3 domains are found, you will be asked to select 3 of them
to search. Both these limits can be overridden as described
the OPTIONS section.

After search scoping is complete, netfind begins a series of
increasingly expensive search phases, after each of which
you will be asked whether to proceed. In phase 1, netfind
attempts to locate two domain name servers that serve each
of the domains isolated during search scoping, and connect
to the SMTP servers on those machines concurrently, request-
ing information about the person being sought. If this is
successful, netfind displays the machine names returned from
the SMTP queries, which will usually include the person's
"home" machine. netfind then attempts to finger the person
at these machines, to provide more information. If this
fails, netfind attempts to finger the person on the domain
name server machines.

In phase 2, netfind searches individual machines found in
the seed database that fall within the domains specified
during search scoping. This phase is more expensive than
phase 1, because dozens of finger searches per domain can be
initiated (as compared with the one to three searches per
domain that execute in phase 1). In phase 3, machines found
while fingering machines in phase 2 are searched. This
final phase does not usually help much, although on occasion
it finds a person who could not be found in the earlier
phases.

Since netfind cannot determine whether a name match
corresponds to the individual actually being sought, when a
match occurs, the currently executing fingers are allowed to
complete, after which time you will be asked if you wish to
continue searching before more fingers are initiated.

Often, once a person is found on one of the concurrently
searched machines, that person will continue to be found on
multiple machines. You should simply interrupt netfind when
this happens.

OPTIONS
    The -h option tells netfind to skip the domain search phase,
    and immediately begin searching individual machines found in
    the seed database. This option exists for measurement

purposes. It is not of much use to casual users.

The −s option will disable usage of the SMTP protocol during
searches. This option is mainly useful for measurement pur-
poses. Without this protocol, searches will begin producing
finger output slightly sooner, but will often search less
useful machines, generate more Internet load, and fail to
find information for users at sites that do not support
finger (such as many companies).

The −t option will report how many timeouts occurred.   The
−T option will set the timeout interval to the specified
number of seconds. It may be necessary to use this option
to increase the timeout value for intercontinental searches.

The −D option sets the maximum number of domains that net-
find will search before asking you to select among domains.
While it may seem convenient to set a high value for this
number, it is suggested that you don't do so. The search
will actually proceed faster (and require less Internet
bandwidth) if a small number of well chosen domains are
searched.

The −H option sets the maximum number of machines that will
be searched by netfind. The default value is 50. Again, we
suggest that you do not set this value higher.

The −m option displays measurement information.   If no
filename is specified, measurements are output to stderr.
The packet count estimates are usually exaggerated, because
they make pessimistic assumptions about the state of the
Domain Naming System (see the paper paper referenced in the
SEE ALSO section below for more details).

The −d option allows you to turn on various classes of
debugging output (all of which are output to stderr), using
a letter corresponding to each one.   Debugging output is
enabled using the −d option with a list of letters, e.g.,
−dsxf. The following classes/letters exist:

a: list the arguments passed to mlook (the main workhorse,
     called after the seed database has been consulted)

c: display control messages (check if the program has
     reached a specified point)

f: display finger related messages

h: list machine names found in the seed database

l: display lock related messages (when entering monitors)

m: display messages about mail protocol (SMTP)

n: display messages about network failures

r: display hosts matched from the seed database that were
      rejected from searches because of search scope selec-
      tion

s: display system call related messages

t: display thread related messages

A: convert the above flags to mean their complement (exam-
      ple: -dAt means produce all debug output except that
      for threads)

The letters that are most likely to be of interest to the
casual user are f, m, and n. By default, these options are
enabled. Specifying any of these three flags with -d on the
command line will disable them (hence, the -d option toggles
the default behavior of each of the flags).

FILES
      /etc/resolv.conf BIND name server locator (needed by resolver)

      RF.* seed database, named according to collection time

      INDEX inverted index into seed database

      SEED_DB environment variable that points to where INDEX is.

BUGS/LIMITATIONS
      1.   Netfind cannot find information about people if neither
            SMTP nor finger is running on some machine in the
            domain being searched, or if the person being sought
            cannot be reached via the Internet.

      2.   Different results can be obtained across invocations
            due to variations in Internet availability.

      3.   Netfind attempts to determine and warn the user if a
            site being searched is not on the Internet. Because
            this is done by checking the names of the authoritative
            Domain name servers for the site, this warning may be
            generated incorrectly in two situations.   First, the
            site may not have Domain servers (e.g., if there are
            not many machines at the site). Second, the site may
            not have its Domain Naming information set up
            correctly. A remedy would be to wait for all search
            attempts to complete before telling the user that a
            site is not on the Internet. However, doing so would
            delay the warning.

4.   SMTP uses a more precise notion of name matching than
     finger. If the specified name does not appear as a
     mail alias, the SMTP portion of a search will fail.

5.   Since netfind uses the SunOS LWP package to support
     inexpensive concurrency (see intro(31)), it can only
     run on machines that support this package, i.e., Suns
     running SunOS 4.0 or more recent.

6.   If SunOS LWP applications exit abnormally, stdin can be
     left in non-blocking input mode, confusing the shell.
     This can manifest itself by the window in which netfind
     was running itself exiting. netfind attempts to handle
     such situations properly, but it may fail to do so at
     times.

7.   Netfind does not recognize some finger responses, since
     there is no standard response format. For example, it
     cannot recognize successful responses from TOPS-20 sys-
     tems.

SEE ALSO
     intro(31), crontab(1), finger(1), invert(1), lookup(1),
     named(8c), nbio(31), ndbm(3), resolver(3), rn(1), send-
     mail(1), yp(1), and "Experience with a Semantically Cog-
     nizant Internet White Pages Directory Tool", by M. F.
     Schwartz and P. G. Tsirigotis, to appear, Journal of Inter-
     networking Research and Experience, 1991. This publication
     discusses the research principles, performance, and scope
     measurements of netfind, and compares it with other white
     pages facilities.

     Netfind is one part of the Networked Resource Discovery Pro-
     ject.   Several   other ideas and prototypes are being
     explored. For more information, contact Schwartz (below).

     Netfind is sold and licensed by XCARET Research, Inc.   This
     version is for a single user only; each user must purchase
     their own copy of Netfind.

     For comments and ordering information, please contact:

          XCARET Research, Inc.
          2060 Broadway, Suite 320
          Boulder, CO 80302
          (303) 444-1285
          netfind@xcaret.com